

PAT-NO: JP411296313A
DOCUMENT-IDENTIFIER: JP 11296313 A
TITLE: STORAGE SUB-SYSTEM
PUBN-DATE: October 29, 1999

INVENTOR-INFORMATION:

NAME	COUNTRY
OBARA, KIYOHIRO	N/A
TAKAMOTO, YOSHIFUMI	N/A

ASSIGNEE-INFORMATION:

NAME	COUNTRY
HITACHI LTD	N/A

APPL-NO: JP10098693
APPL-DATE: April 10, 1998

INT-CL (IPC): G06F003/06 , G06F013/00

ABSTRACT:

PROBLEM TO BE SOLVED: To attain a load distribution between controllers and to re-arrange a managed file and a volume by connecting plural disk controllers to plural disk driving devices by means of a network or a switch.

SOLUTION: A route change processor in a route controller 21 refers to the transmission destination control unit number and the volume number of a frame which is transmitted from a host and transmits the frame to the disk controller which actually controls the volume to be the object of the frame in accordance with a volume-disk controller correspondence converting table 29. When accesses are concentrated on the volumes A11 and A12, for example, the load of the disk controller 8a becomes heavy and the loads of the controllers 8a and 8b are unbalanced. In this case, the table 29 is changed, the accesses to the volumes A13 and A14 are enabled to be transmitted to the disk controller (b) and, moreover, the disk controller 8b is made to be in charge of control for the volumes A13 and A14.

COPYRIGHT: (C)1999,JPO

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-296313

(43) 公開日 平成11年(1999)10月29日

(51) Int.Cl.⁸G 0 6 F 3/06
13/00

識別記号

3 0 4
3 5 7

F I

G 0 6 F 3/06
13/003 0 4 B
3 5 7 C

審査請求 未請求 請求項の数18 O L (全 15 頁)

(21) 出願番号 特願平10-98693

(22) 出願日 平成10年(1998)4月10日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 小原 清弘

東京都国分寺市東恋ヶ窪一丁目280番地

株式会社日立製作所中央研究所内

(72) 発明者 高本 良史

東京都国分寺市東恋ヶ窪一丁目280番地

株式会社日立製作所中央研究所内

(74) 代理人 弁理士 小川 勝男

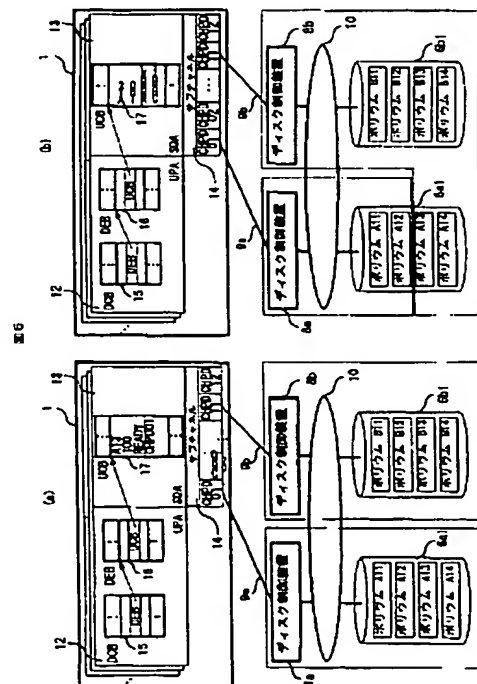
(54) 【発明の名称】 記憶サブシステム

(57) 【要約】

【課題】 ホスト上の適用業務やディスク制御装置は通常の処理を続けたまま、ビジーなディスク制御装置下の任意のボリュームの管理を、ビジーでない任意の別のディスク制御装置へ移行する手段を提供する。

【解決手段】 複数のディスク制御装置と複数のディスク駆動装置とをネットワーク又はスイッチで接続する。また、ディスク制御装置の負荷に応じてディスク制御装置間で管理するボリュームを交代させる手段と、ボリュームの管理元の交代に対応してホストからディスク制御装置へのアクセスパスを変更する手段と、ホストコンピュータが指定したボリューム番号とアクセスパスとの対応を交換する手段を設ける。

【効果】 ボリュームの交代やアクセスパスの変更をホストプロセッサから隠蔽して行うことができるので、複数のディスク制御装置間でのボリューム再配置による負荷分散を適用プログラムを停止せずに行うことができる。



【特許請求の範囲】

【請求項1】複数のディスク制御装置と複数のディスク駆動装置とを有し、

上記複数のディスク制御装置と上記複数のディスク駆動装置とをネットワークもしくはスイッチにより接続したこと特徴とする記憶サブシステム。

【請求項2】上記複数のディスク駆動装置の一のディスク制御装置の管理元を、上記複数のディスク制御装置の一のディスク制御装置から他のディスク制御装置に変更する管理元変更手段を有することを特徴とする請求項1記載の記憶サブシステム。

【請求項3】上記管理元変更手段が行う管理元の変更は、該記憶サブシステムに接続されたホストシステム上で上記記憶サブシステムと入出力を行っている適用業務プログラムに対し隠蔽されて行われることを特徴とする請求項2記載の記憶サブシステム。

【請求項4】上記ディスク制御装置と上記ディスク駆動装置との一組でも、独立した記憶サブシステムとして動作し得ることを特徴とする請求項2記載の記憶サブシステム。

【請求項5】上記複数のディスク制御装置間で通信を行うための通信手段を有することを特徴とする請求項1記載の記憶サブシステム。

【請求項6】上記管理元変更手段は、ディスク制御装置を増設した場合に、上記複数のディスク駆動装置の位置のディスク駆動装置の管理元を、上記増設したディスク制御装置に変更することを特徴とする請求項2記載の記憶サブシステム。

【請求項7】上記複数のディスク制御装置の負荷をモニタするモニタ手段を有し、
上記管理元変更手段は、上記モニタ手段のモニタ結果に応じて、管理元の変更を行うことを特徴とする請求項2記載の記憶サブシステム。

【請求項8】上記モニタ手段によりモニタされる負荷は、上記複数のディスク制御装置の各ディスク制御装置内、及び上記複数のディスク制御装置間を接続する装置内のデバイスの利用率であることを特徴とする請求項7記載の記憶サブシステム。

【請求項9】上記デバイスは、プロセッサ、バス、又はメモリであることを特徴とする請求項8記載の記憶サブシステム。

【請求項10】上記管理元変更手段が行う変更は、ホストシステム又は上記複数のディスク制御装置の何れかのディスク制御装置からの指示に応じて行われることを特徴とする請求項2記載の記憶サブシステム。

【請求項11】上記変更手段を管理する制御装置を有し、
上記変更手段が行う変更は上記制御装置の指示により行われることを特徴とする請求項2記載の記憶サブシステム。

【請求項12】ホストシステムが管理する、上記ホストシステムと上記複数のディスク制御装置との接続経路情報を保持する手段を有し、

上記接続経路情報を保持する手段は、上記管理元変更手段が行う管理元の変更に応じて、上記接続経路情報を変更することを特徴とする請求項2記載の記憶サブシステム。

【請求項13】ホストシステムと記憶サブシステムとの間に、データ交換の経路を変更する経路変更装置を備え、

上記経路変更装置は、上記管理元変更手段が行う管理元の変更に応じて、上記データ交換の経路を変更することを特徴とする請求項2記載の記憶サブシステム。

【請求項14】該ディスク駆動装置または該ディスク駆動装置内のデバイスと、それらの制御を受け持つ該ディスク制御装置との対応の変更と同期して、ホストシステム上に存在する、ホストシステムと該複数のディスク制御装置との接続経路情報の変更であって、

該ディスク駆動装置または該ディスク駆動装置内のデバイスに対する接続経路情報が、該ディスク駆動装置または該ディスク駆動装置を制御する該ディスク駆動装置になるように変更することを特徴とする請求項7記載の記憶サブシステム。

【請求項15】上記接続経路情報の変更、及び上記管理元変更手段が行う管理元の変更が終了するまで、上記ホストシステムからの該記憶サブシステムへの入出力を一時中断されることを特徴とする請求項12記載の記憶サブシステム。

【請求項16】複数のディスク制御装置と、
上記複数のディスク制御装置とネットワーク又はスイッチで接続された複数のディスク駆動装置と、
上記複数のディスク制御装置の各ディスク制御装置が管理する上記複数のディスク駆動装置の各ディスク駆動装置内のボリュームを管理する管理元を、上記複数のディスク制御装置の一のディスク制御装置を他のディスク制御装置に変更する管理元変更手段と、
上記変更に対応してホストからディスク制御装置へのアクセスパスを変更するアクセスパス変更手段、とを有することを特徴とする記憶サブシステム。

【請求項17】ホストコンピュータにより指定されたボリューム番号と上記アクセスパスとの対応を保持する手段を有することを特徴とする請求項16に記載の記憶サブシステム。

【請求項18】上記管理元変更手段は、上記複数のディスク制御装置の負荷に応じて上記管理元の変更を行うことを特徴とする請求項16又は請求項17の何れかに記載の記憶サブシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、複数のディスク制

御装置間又はディスク制御装置内の負荷分散に関する。特にディスク制御装置とディスク駆動装置を、ネットワークもしくはスイッチ等で結合したディスクシステムでの負荷分散に関する。

【0002】

【従来の技術】既設置のディスクシステムの性能や容量を増加させるには、新たなディスク制御装置とそれに付随するディスク駆動装置を追加すればよい。すなわち、性能を求める場合はディスク制御装置当たりのディスク駆動装置を少なく設定し、ディスク制御装置の負荷を少なくすればよい。また、容量を求める場合は、ディスク制御装置当たりのディスク駆動装置を多く設定すればよい。このような、複数のディスク制御装置とディスク駆動装置を並べる構成により、性能や容量に対して広いスケラビリティを提供できるように見える。

【0003】しかし、この形態におけるスケラビリティの広さは、全てのディスク制御装置やディスク駆動装置内のボリュームへのアクセスが、均等に起きることを前提としている。実際の運用ではこれらのアクセスには偏りがあるため、ディスク制御装置およびボリュームの負荷は偏った状態となる。この偏りは、ホストコンピュータ上で動作する適用業務の性質に依存し、また、同一適用業務においても時間によって変動する。この例としては、昼間のオンライントランザクション処理と夜間のバッチ処理の関係が挙げられる。すなわちディスク制御装置を複数個並べた構成では、ディスク制御装置間で負荷の不均衡が起り、設置されている全部の制御装置の性能を十分に引き出すことは困難である。

【0004】従来、この不均衡への対応として、ユーザもしくはシステムエンジニアが、ディスク制御装置間でファイルやボリュームを再配置するなどの調整を行っていた。すなわち、人手による調整であった。このため、計算機の稼働状況の変化に合わせた素早い対応は不可能であった。また、再配置処理自体も、ホストプロセッサ経由で多量のデータ転送を行うため、時間がかかる非常に重たい処理であり、ホストの負荷が増加する等の不都合が多かった。さらに、データの整合性を保つため再配置処理中はデータの更新は禁止されていた。このように、再配置中のデータを利用する適用業務は、再配置が終了するまで実行できないなど運用上も非常に不都合が多い処理であったため、複数のディスク制御装置を並べた従来型のディスクシステムでは負荷の不均衡の発生は必然であり、これら複数のディスク制御装置全てが同時に最大性能を発揮することは実質的に不可能であった。

【0005】これに対し、単一のディスク制御装置は、ホストコンピュータとのインタフェースを行うチャンネルインタフェース、ディスク駆動装置とのインタフェースを行うディスクインタフェースおよびキャッシュメモリなどのコンポーネントにより構成されている。これら単一ディスク制御装置内のコンポーネントは、相互接続性

の高い共通バス等の手段により結合されており、チャンネルインタフェースやディスクインタフェースは複数個実装されているのが一般的である

単一のディスク制御装置は、上記のような構造上の理由により、装置内の負荷分散は容易に行える。すなわち、アクセスが集中するディスク駆動装置に対し、チャンネルインタフェースやキャッシュ等の資源を優先して割り当てることにより、ディスク制御装置全体として最良のパフォーマンスを提供できる。つまり同一容量のディスクシステムを構築する場合、複数のディスク制御装置を用いるよりも、単一のディスク制御装置を用いた方が、アクセスの偏りによる負荷の不均衡に強く、前記ファイルやボリュームの再配置処理も不要になる。

【0006】ところが近年、単体のディスク駆動装置の容量の増大や、ディスクアレイ技術の発達により、単一のディスク制御装置が管理するディスク駆動装置およびチャンネルバス（ホストとのバス）の数が非常に増加してきている。このため現在の技術では、単体のディスク制御装置としての構成を保ったまま、管理するディスク駆動装置やチャンネルバスを飛躍的に増加させることは困難である。

【0007】この理由は次のとおりである。記憶容量およびチャンネルバスの増加により、ディスクインタフェースやチャンネルインタフェースのコンポーネント数の増加する。このためディスク制御装置の内部において処理が集中するボトルネックを生み出し、単一のディスク制御装置の性能向上を妨げている。このようなボトルネックの例としては、共通バスが挙げられる。すなわち高い相互接続性を提供するため、全てのコンポーネント間通信が共通バスを経由するからである。

【0008】つまり単一のディスク制御装置の利点の一つである、共通バス等によるコンポーネント間の柔軟性の高い結合が、一方ではボトルネックとなり、単一のディスク制御装置全体の処理能力が律速されることとなる。

【0009】

【発明が解決しようとする課題】このようなボトルネックによる性能の律速の問題を回避し、ディスク制御装置の性能をスケラブルに向上させるには、複数のディスク制御装置とディスク駆動装置を並べる構成が本質的に適している。すなわち、共有部分を少なくする構成である。

【0010】しかし、前述したように、従来のかかる構成の装置では、ディスク制御装置間の負荷分散を効率よく行うことができなかった。すなわち、ディスク制御装置間での負荷の不均衡という問題を効率よく解消することができなかった。

【0011】そこで、本願発明者は、ディスク制御装置間での負荷の不均衡という問題の本質を検討した結果、以下の点が明らかとなった。

【0012】上記の複数のディスク制御装置とディスク駆動装置を並べた場合の負荷の不均衡の問題とは、次のような状態である。あるディスク制御装置に接続されている複数のディスク駆動装置又はボリウムの中で、一部のディスク駆動装置又はボリウムにホストからのアクセスが集中した場合、それらを管理するディスク制御装置の負荷が高まり過負荷になる。ここでは、ディスク制御装置の性能限界は、その配下のディスク駆動装置の性能限界より低い場合を想定している。一般に、一つのディスク制御装置は多くのディスク駆動装置を管理するので、例えばあるディスク制御装置が10個のディスク駆動装置を管理している場合、その中の5個のディスク駆動装置にアクセスが集中すると、ディスク制御装置の方が先に過負荷になるのが一般的だからである。

【0013】従来の装置では、ディスク駆動装置やボリウムと、そのボリウムを管理するディスク制御装置との対応が固定されているため、アクセスが集中していないディスク駆動装置又はボリウムに対してアクセスする場合でも、過負荷なディスク制御装置を経由しなければならない場合が生じる。この場合、目的のディスク駆動装置又はボリウム自体の負荷は低いにも係わらず、これらに対するアクセスは、過負荷なディスク制御装置を経由するため待たされていた。

【0014】上述の検討により、ディスク制御装置の過負荷は、当該ディスク制御装置が管理する一部のディスク駆動装置又はボリウムにホストからのアクセスが集中することに問題があり、当該アクセス集中の問題は、ディスク駆動装置やボリウムと、そのボリウムを管理するディスク制御装置との対応が固定されていることに問題があることが判明した。

【0015】そこで、本発明の目的は、前記のアクセス集中時において、ホスト上の適用業務やディスク制御装置は通常の処理を続けたまま、過負荷なディスク制御装置下の任意のディスク制御装置又はボリウムの管理を、過負荷でない任意の別のディスク制御装置へ移行する手段を提供することにある。これにより、複数のディスク制御装置間で負荷分散を実現し、ディスクシステム全体として、これら複数のディスク制御装置全てが同時に最大性能を発揮できる、コストパフォーマンスの良いディスクシステムを提供する。別の見方をすれば、本発明の目的は、実際のデータ移動を行わずに、任意のディスク制御装置間で、管理されるファイルやボリウムの再配置が可能なディスクシステムを提供することにある。さらに、本発明の別の目的は、ディスク制御装置を増設した場合に、自動的にファイルやボリウムの再配置を行い、複数のディスク制御装置間で自律的に負荷をバランスさせるディスクシステムを提供することにある。また、この再配置や負荷のバランスを、ディスク制御装置だけを追加した場合にも適用可能とすることにある。さらに、本発明の別の目的は、前述のボリウムの管理の移行や再

配置による変更を、ホストプロセッサから隠蔽が可能なディスクシステムを提供することにある。

【0016】なお、シェアドナッシング方式のデータベース管理システムが、疎結合プロセッサシステム上で動いているシステムに対し、各プロセッサの負荷が均等になるように、各プロセッサと、各プロセッサがアクセスする外部記憶装置との接続を動的に変更する旨が、特開平9-73411号公報（以下「文献1」という。）に記載されている。しかし、文献1には、ホストの負荷分散を開示するものであり、ディスク制御装置の負荷分散、及びファイル再配置時のホストの負荷軽減については開示されていない。

【0017】また、アクセス頻度に応じてファイルを装置間でマイグレーションする旨が特開平3-294938号公報（以下「文献2」という。）に開示されている。しかし、文献2には、ファイルを装置間で移動せずにファイルの再配置を行う点は開示されていない。

【0018】

【課題を解決するための手段】上記目的を解決するために、本発明では、複数のディスク制御装置と複数のディスク駆動装置とをネットワーク又はスイッチで接続する。また、各ディスク制御装置の負荷に応じてディスク制御装置間で管理するボリウムを交代させる手段と、前記ボリウム交代に対応してホストからディスク制御装置へのアクセスパスを変更する手段と、ボリウムの交代やアクセスパスの変更をホストプロセッサから隠蔽するため、ホストコンピュータが指定したボリウム番号とアクセスパスを変換する手段を設ける。

【0019】

【発明の実施の形態】図1～図12を用いて本発明の一実施例を説明する。

【0020】図2に単一のディスク制御装置の一構成例を示す。ディスク制御装置8は、ホストインタフェース2とチャネルバス9によりホストコンピュータ1と接続し、共有メモリ3、キャッシュメモリ4、複数のディスク駆動装置6と接続するディスクインタフェース5およびこれらを接続する共通バス7により構成されている。複数のディスクインタフェースが実装可能である。複数のホストインタフェース2も実装可能であり、この場合接続先のホスト1が同一であるか否かは問わない。本実施例の図2では、各ホストインタフェースが同じホストに接続されている例を示している。

【0021】ホストインタフェース2とディスクインタフェース5には、プロセッサが装備されており、それぞれ自律的に動作する。また、共有メモリ3とキャッシュメモリ4は、複数のホストインタフェース2や複数のディスクインタフェース5から参照可能な共有資源である。キャッシュメモリ4には、本ディスク制御装置へ書き込まれたデータやディスクドライブ6から読み出されホストに出力したデータが一時的に保存される。

【0022】共有メモリ3には、キャッシュメモリ上のデータの管理情報が格納される。この情報をキャッシュディレクトリと呼ぶ。ホストインタフェース2やディスクインタフェース5は、ホストから与えられるリード/ライトコマンド等に記述されている、ボリウム番号やボリウム内のレコードの位置情報(シリンダ番号、ヘッド番号、レコード番号等)をキーとして、キャッシュディレクトリの情報を検索することにより、リード/ライトの対象としているデータが、キャッシュメモリ上にある(キャッシュヒット)かキャッシュメモリ上にない(キャッシュミス)かの判断、およびキャッシュヒットの場合はそのデータのキャッシュメモリ上の位置を取得できる。共有メモリ3とキャッシュメモリ4は、ハードウェア的に同じ構成で実現できるが、ここでは格納される内容によって、共有メモリとキャッシュメモリという区別を行う。すなわち、純粋な入出力データがキャッシュメモリ上にあり、キャッシュメモリ上のデータ管理に関する制御情報は共有メモリにある。

【0023】また共有メモリ上には、サブシステム構成情報として、接続されているディスク駆動装置やチャネルパスの数、キャッシュメモリや共有メモリの容量等、ディスク制御装置の構成に関する詳細情報が格納される。さらに、ボリウム情報として、ボリウムが格納されているディスク駆動装置や接続されているディスクインタフェースに関する情報、ボリウムの容量やフォーマット等の情報が置かれている。

【0024】ディスク制御装置がディスクアレイ機能を持っている場合には、ホストから送られたデータは分割され、複数のディスク駆動装置や、ディスク駆動装置を構成する複数のHDA(ヘッド・ディスク・アセンブリ)に分散して格納される。このように、ディスクアレイ構成を採用しデータを分散して格納している場合、ボリウム情報には、ディスクアレイの構成情報として、そのボリウムが利用しているディスク駆動装置やHDAのリストも格納される。本実施例はディスクアレイに対しても適用可能であるが、説明の簡略化のため、以後の説明は通常のディスクに対する動作を例に説明する。

【0025】次に、ホストコンピュータ1とディスク制御装置8間でのデータの入出力の動作について説明する。

【0026】ホストコンピュータ1とディスク制御装置8の間では、チャネルパス9を通して、一定の形式でデータ交換すなわち入出力が行われる。このデータをフレームと呼ぶ。フレームのフォーマットを図10に示す。送信先コントロールユニット番号フィールド22と発信元コントロールユニット番号フィールド23は、それぞれ送信元と発信先のコントロールユニット番号を示す。コントロールユニット番号とは、ディスク制御装置を識別するために、各ディスク制御装置毎に固有に付される番号である。ホストコンピュータに対するコントロール

ユニット番号は0である。

【0027】ボリウム番号フィールド24は、送信先コントロールユニット番号フィールド22や発信元コントロールユニット番号フィールド23で指定されたディスク制御装置内のボリウムを指定する情報であり、ここで指定されたボリウムに対し、後に続くコマンド/データフィールドの指示や情報が適用される。コマンド/データフィールド25は、コマンドやデータが入るフィールドである。一般に、その他にもいくつかのフィールドが存在するが、本実施例においては参照しないため省略する。

【0028】ここで、ホストコンピュータ2上の適用プログラムが、コントロールユニット番号8番のディスク制御装置内の10番のボリウムから、シリンダ番号2番、ヘッド番号3番、レコード番号4番のデータを読む命令を実行した場合、ホストコンピュータ2からディスク制御装置8に送信されるフレームの内容は、送信先コントロールユニット番号フィールドには8が、発信元コントロールユニット番号フィールドには0が、ボリウム番号フィールド24には10が入り、コマンド/データフィールド25には、リードコマンドとそのパラメタである、シリンダ番号2番、ヘッド番号3番、レコード番号4番が入る。

【0029】フレームはホストインタフェース2により受信される。ホストインタフェース2では、ボリウム番号フィールド24およびコマンド/データフィールドの情報により、操作の対象となるボリウムと操作の種類を認識する。また、共有メモリ上のボリウム情報を参照することにより、このボリウムの操作を担当するディスクインタフェースを認識し、そのディスクインタフェースに処理を委託する。

【0030】ボリウム情報の一例を図3に示す。ここで、ボリウム情報は表形式になっており、ボリウム番号をキーとして、そのボリウムに関する情報を取得できる。ボリウム容量やフォーマットに関する情報は、本実施例では参照しないため省略する。

【0031】処理を委託されたディスクインタフェースもまた、ボリウム情報を参照し、このボリウムが格納されているディスク駆動装置番号、HDA番号、シリンダオフセットを取得する。この情報に従い、ディスクインタフェースは目的のボリウムを格納しているディスク駆動装置のHDAに入出力命令を発行する。前に説明したフレームに対する場合、ボリウム番号は10番のため、図3のボリウム情報より、1番のディスク駆動装置の、0番のHDAに対し、シリンダ番号2番、ヘッド番号3番、レコード番号4番に対応するデータを読むように、リードコマンドを発行する。

【0032】ここで一つのHDAの容量は、一つのボリウム容量より大きい。このため、一つのHDAに複数のボリウムが格納されている。これら一つのHDA中のボ

リウムの切り分けはシリンダオフセットにより行われる。すなわち、図3のポリウム情報の例では、0番のHDA中の、シリンダ0番からシリンダ19番までがポリウム番号10の範囲であり、シリンダ20番からがポリウム番号11の範囲である。このため、ポリウム番号11の場合のような、実際のシリンダが0番から始まっていないポリウムに対しては、ディスクインタフェースがHDAにコマンドを発行するときに、シリンダ番号に対しオフセットの値を加え、ホストから見たシリンダ番号からHDA中の実際のシリンダ番号への変換を行う。

【0033】入出力動作の説明に戻る。HDAはリードコマンドに対応して読み出されたデータをディスクインタフェースに返す。ディスクインタフェースはそのデータを、ホストインタフェースに返す。ホストインタフェースは、読み出されたデータに対し新たなフレームを生成して、ホストコンピュータ1に送信する。このフレームの内容は、さきほどホストコンピュータ1から送られてきたフレームの内容と比較して、送信先コントロールユニット番号フィールドと、発信元コントロールユニット番号フィールドの内容が入れ替わり、コマンド/データフィールドの内容が要求された読みだしデータに置き換えられる。

【0034】以上のようにして、ホストコンピュータ1とディスク制御装置8間でデータの入出力が行われる。

【0035】図1に、本発明によるディスク制御装置とディスク駆動装置の一構成例を示す。本構成では、ディスク制御装置8aから8cとディスク駆動装置6a1から6c2の間をデバイスネットワーク10で結合している。デバイスネットワーク10は、各ディスク駆動装置が、全てのディスク制御装置およびその中のポリウムを

直接アクセス可能な構成を実現する。

【0036】デバイスネットワーク10は、一般的なSCSIバスで構成することも可能であるが、より多くのディスク制御装置及びディスク駆動装置を接続するため、ファイバチャネル等のネットワークの利用が適している。また、デバイスネットワークの総合的なスループットを高めるため、スイッチで構成する手法も適している。また同様にスループットを高めるため、複数系統のネットワークやスイッチを用いることもできる。

【0037】各ディスク制御装置には、初期状態として、各ディスク制御装置が制御を担当するディスク駆動装置およびディスク駆動装置内のポリウムが割り当てられている。例えば図1では、ディスク駆動装置8aはディスク駆動装置6a1および6a2内のポリウム、ディスク駆動装置8bはディスク駆動装置6b1および6b2内のポリウム、ディスク駆動装置8cはディスク駆動装置6c1および6c2内のポリウムの制御を、それぞれ担当する。

【0038】ディスク制御装置と、そのディスク制御装置が制御するディスク駆動装置やポリウムとの対応は、

ホストコンピュータ上のオペレーティングシステムにも記憶されている。したがって一般には、ホストコンピュータを稼働させたまま、ディスク制御装置と、そのディスク制御装置が制御するポリウムとの対応を変更することは出来ない。

【0039】これに対し、ホストコンピュータを稼働させたまま、ディスク制御装置と、そのディスク制御装置が制御するポリウムとの対応の変更を実現するため、ホストコンピュータとディスク制御装置間でのフレームの交換の制御を行う機構が、経路制御装置21である。

【0040】経路制御装置21の一構成例を図11に示す。ホストチャネルインタフェース26は、ホストコンピュータとのインタフェースを行う。ディスクチャネルインタフェース27は、ディスク制御装置とのインタフェースを行う。ホストコンピュータ側とディスク制御装置側とで同一のプロトコルを用いる場合は、ホストチャネルインタフェース26とディスクチャネルインタフェース27は同一のハードウェア構成をとることが出来る。

【0041】経路変更プロセッサ28は、内部にプロセッサとメモリおよび制御プログラムを持つ。この制御プログラムの動作は、次のとおりである。ホストチャネルインタフェース26を通して送られたフレームを、適切なディスク制御装置が接続されているディスクチャネルインタフェース27に送り、ディスクチャネルインタフェース27はそのフレームをディスク制御装置に送信する。また逆に、ディスク制御装置からのフレームをホストコンピュータに返す。ここで、適切なディスク制御装置とは、そのフレームの送信先コントロールユニット番号とポリウム番号を、ポリウム-ディスク制御装置対応変換表29に基づいて変換を行った後のディスク制御装置を指す。

【0042】ポリウム-ディスク制御装置対応変換表29は、オペレーティングシステムが記憶しているポリウムとディスク制御装置の対応から、経路制御装置21の配下にある、ポリウムとディスク制御装置の対応を導く表である。この表は、オペレーティングシステムが記憶しているポリウムとディスク制御装置の対応とは独立しているため、ホストコンピュータを稼働させたまま、オペレーティングシステムやホスト上の適用プログラムを変更せずに、任意のディスク制御装置に任意のポリウムに対するフレームを送信することが可能である。すなわち、ディスク制御装置と、そのディスク制御装置が制御するポリウムとの対応の変更を、ホストコンピュータに対し隠蔽できる。

【0043】ポリウム再配置制御プロセッサ30は、内部にマイクロプロセッサを持ち、各ディスク制御装置の状態に基づき、ポリウム-ディスク制御装置対応変換表29を変更することにより、経路変更プロセッサによるフレームの送信先を制御する。ディスク制御装置の状態

11

に基づく具体的な送信先の制御方法は後述する。

【0044】各ディスク制御装置の状態は、デバイスネットワークインタフェース31を通して収集する。すなわち本構成例では、ディスク制御装置間および経路制御装置で、ホストコンピュータを介さず情報や入出力データの交換を行うことが可能である。つまりデバイスネットワーク10は、ディスク駆動装置内に格納されているデータを転送するだけでなく、ディスク制御装置間の情報交換ネットワークとしても用いる。もちろん、ホストコンピュータを介したり、ディスク制御装置間にデバイスネットワークとは別のネットワークを構築し、それらを用いる構成でも同様の情報交換は可能である。

【0045】次に本発明における負荷分散の実現方式を説明する。本発明の負荷分散は、ディスク制御装置と、そのディスク制御装置に制御されるボリウムの対応を任意に変更し、この変更をホストコンピュータやホストコンピュータ上の適用プログラムから隠蔽することにより実現する。これは、ディスク制御装置間でのデータの移動を伴わないボリウムの再配置と、その再配置のホストコンピュータや適用プログラムへの隠蔽とも考えられ

る。

【0046】例えば初期状態として、図4(a)の細線で示すように、ディスク制御装置8aがボリウムA11からA14までを制御し、ディスク制御装置8bがボリウムB11からB14までを制御していたとする。このようなディスク制御装置とボリウムの対応は、経路制御装置21内のボリウム-ディスク制御装置対応変換表29に反映されている。

【0047】図4のボリウム-ディスク制御装置対応変換表は、各行が、ボリウムとディスク制御装置の対応を表す。最初と次の列が、ホストコンピュータから見たボリウムとそのボリウムを制御するディスク制御装置のコントロールユニット番号、最後の列が実際に前記ボリウムを制御するディスク制御装置のコントロールユニット番号を表す。例えば最初の行は、ホストコンピュータから、ディスク制御装置8aが制御するボリウムA11として認識されているボリウムが、実際にも、ディスク制御装置8aが制御していることを示す。経路制御装置21内の経路変更プロセッサ28は、ホストから送られたフレームの、送信先コントロールユニット番号とボリウム番号を参照し、このボリウム-ディスク制御装置対応変換表に従い、実際にそのフレームが対象とするボリウムを制御するディスク制御装置にフレームを送信する。

【0048】さてここで、ボリウムA11とA12にアクセスが集中した場合を考える。この場合、ディスク制御装置8aの負荷が重くなり、ディスク制御装置8aと8bの負荷が不均衡になる。また、ボリウムA13やA14に対するアクセスも、負荷の重いディスク制御装置8aを経由する必要があると待たされる。これに対し、図

12

4(b)で示すように、ボリウム-ディスク制御装置対応変換表29を変更し、ボリウムA13とボリウムA14へのアクセスが、ディスク制御装置bに送られるようにし、さらにディスク制御装置8bに、ボリウムA13とA14の制御を担当させる。

【0049】以上のようにして、ホストコンピュータおよびその上で動作する適用プログラムには変更を加えず、また適用プログラムは動作を継続させたまま、ディスク制御装置間でボリウムの再配置を行い、ディスク制御装置8aの負荷を軽減し、システム全体として最良のスループットを提供できる。

【0050】次に、ディスク制御装置間でのデータの移動を伴わないボリウムの再配置による負荷分散の、具体的な手続きを図5のフローチャートを用いて説明する。

【0051】各ディスク制御装置はそれぞれの資源に対する利用率をモニタし負荷を判定する。資源の例としては、チャネルバス、ホストインタフェース、共有メモリ、キャッシュメモリ、共通バス、ディスクインタフェース、ディスク駆動装置とのバス等が挙げられる。ディスク制御装置は、これらの資源の利用率を、自分の制御下にあるボリウム毎の負荷としてモニタする。利用率の具体例としては、ホストインタフェース、ディスクインタフェースに関しては制御を行うプロセッサの稼働率、共有メモリ、キャッシュメモリ等に関しては単位時間あたりのアクセス回数、共通バス、チャネルバス、ディスク駆動装置とのバス等に関しては単位時間あたりに利用されている割合等が挙げられる。

【0052】各ディスク制御装置でモニタされたボリウム毎のディスク制御装置の負荷の情報は、デバイスネットワーク10等の情報交換網を通じてボリウム再配置制御プロセッサ30に集められる。ボリウム再配置制御プロセッサ30では、この情報を元に各ディスク制御装置の負荷を計算し、限度以上の負荷の不均衡が生じている場合は、各ディスク制御装置の負荷が均等になるようなボリウムの配置を決定する。この限度は、あらかじめ利用者が指定しておく。具体的な配置は、自動最適配置が指定されている場合は、ボリウム再配置制御プロセッサ30が各ディスク制御装置の負荷が均等になるように自律的に決定する。それ以外の場合は、ユーザが決定し指示する。

【0053】次にボリウム再配置制御プロセッサ30は、経路変更プロセッサ28の機能を一時保留させることにより、ホストコンピュータ1とディスク制御装置8間の入出力を一時保留させる。これはボリウムの再配置中の過渡状態での入出力を抑制するためである。

【0054】そしてボリウム再配置制御プロセッサ30は、決定したボリウムの配置に基づき、ボリウム-ディスク制御装置対応変換表29を変更するとともに、決定したボリウムの配置を、各ディスク制御装置に指示する。各ディスク制御装置は指示に従い、管理するボリウ

ムを変更する。デバイスネットワークにより任意のディスク制御装置が任意のディスク駆動装置を直接アクセスできるため、この変更による物理的な接続の変更は不要である。

【0055】あるディスク制御装置がこれまで制御していたボリュームの管理を止める場合は、そのボリュームに対するアクセス要求を拒否するだけでよい。

【0056】一方、あるディスク制御装置が他のディスク制御装置が管理していたボリュームを引き継いで管理するには、対象となるボリュームのボリューム情報が必要になる。具体的には図3に示した各ディスク制御装置の共有メモリ上に格納されている情報であり、そのボリュームが格納されているディスク駆動装置のデバイスネットワーク上での位置（ディスク駆動装置番号）や、ボリュームの容量やフォーマット等の情報である。このボリューム情報は、ボリュームの再配置の実行に先立ち、移動元と異動先のディスク制御装置の間で交換しておく。この交換もデバイスネットワークを用いる。デバイスネットワークではなく、専用のネットワークを用いる方式を採用することももちろん可能である。また、全ディスク制御装置の中に、事前に、その制御装置がアクセスできるディスク駆動装置やボリュームに対する情報等をすべて格納しておけば、ボリュームの再配置に先だってボリュームに関する情報交換を行う必要はない。

【0057】最後に、ボリューム再配置制御プロセッサ30は、保留していた経路変更プロセッサ28の機能を回復させることにより入出力を再開させる。以上の手順により、ボリュームの再配置による負荷分散が行われる。

【0058】図4では、ディスク制御装置とディスク駆動装置がそれぞれ2台の場合を例に取ったが、本発明の負荷分散は、デバイスネットワークで結合した任意台数のディスク制御装置と任意台数のディスク駆動装置及びボリュームに対して適用できる。また、ディスク制御装置の能力が足りない場合は、ディスク駆動装置の台数はそのまま、ディスク制御装置だけを増設し負荷分散することにより、システム全体のスループットを高めることもできる。

【0059】本発明の特徴の一つは、ボリューム再配置を行っても実際のデータの移動が不要な点である。このため、秒オーダー以下の非常に短い時間での再配置が可能となる。また、その他の特徴として、ディスク制御装置間で共有し同時に参照し合う情報が無い事も挙げられる。すなわちボリューム再配置時以外は、ディスク制御装置とその配下のボリュームは、従来の単体のディスク制御装置と同様の動作を行う。特に、他のディスク制御装置との間で、ボリュームに格納する（されていた）入出力データのやりとりを行わない。これにより、ディスク制御装置間のスムーズな負荷分散の実現と、ディスク制御装置の増設によるスケーラブルな性能向上の両立が達成できる。

【0060】本発明の目的は、経路制御装置を使用しなくても達成可能である。この方式を第二の実施例として次に示す。

【0061】図9に第二の実施例におけるディスク制御装置の一構成例を示す。本実施例においては、経路制御装置を使用しない。このため、例えば図12の(a)から(b)のように、各ディスク制御装置が制御するボリュームが変更された場合、ボリュームA13に対するフレームは、変更後もディスク制御装置8aに届いてしまう。ここで各ディスク制御装置が担当するボリュームの範囲を細線で示している。

【0062】このような誤配を防ぎ、かつディスク制御装置とその制御装置に制御されるボリュームの変更を、ホストコンピュータ上の適用プログラムから隠蔽するために、本実施例では、ホストコンピュータ1上で動作する入出力制御ユーティリティプログラムを用いる。

【0063】入出力制御ユーティリティを用いた、ディスク制御装置と、そのディスク制御装置に制御されるボリュームの変更の具体的な隠蔽方式を次に述べる。

【0064】ここでは、図12の(a)から(b)のように、各ディスク制御装置が制御するボリュームを変更する場合について述べる。図6(a)は、図12(a)の状態の時点の、ホストコンピュータ1内部のディスクアクセスに関する制御構造を詳細に示した図である。ホストのオペレーティングシステムは、IBM社により開発されたMVS（マルチプル・バーチャル・ストレージ）と仮定する。この図において、UPA12は、ユーザプログラムエリアの略であり、仮想記憶アドレス内のユーザ領域を示す。同様にSQA13は、システム・キュー・エリアの略であり、仮想記憶アドレス内のオペレーティングシステムの領域を示す。サブチャネル14は、実際にディスク制御装置等との入出力を実行する機構であり、チャンネルバス9a、9bが接続されている。この図のシステムでは、最大12本のチャンネルを接続可能であり、各接続機構毎にCHPD01からCHPD12までのIDが割り当てられている。

【0065】この図を用い、ボリュームをアクセスする従来の手順を説明する。この図で示したUPA12は、ディスク制御装置8aを用い、ボリュームA13をアクセスする。具体的な手順は、次の通りである。

【0066】ディスクアクセスに当たり、最初に参照する構造体は、自分のUPA内のDCB（データ・コントロール・ブロック）15である。DCBは、アクセス対象となるデータセットをオープンしたときに作成され、データセットの構成、レコードの論理長、レコードフォーマットの情報等に加え、DEB（データ・エクステンション・ブロック）16へのポインタが格納されている。

【0067】DEBもアクセス対象となるデータセットをオープンしたときに作成され、データセットの詳細情

報、すなわちアクセス方式やデータセットがDASD上に占めている物理的な区域の情報等に加え、UCB(ユニット・コントロール・ブロック)17へのポイントが格納されている。

【0068】UCBは、システム・キュー・エリアにある情報であり、ホストシステムの立ち上げ時に、ハードウェア構成定義で設定された情報に基づき、それぞれのボリュームに対し生成される。また、内容は後述する入出力制御ユーティリティにより変更できる。UCBには、ボリュームの詳細情報、すなわちボリューム名、ボリュームシリアル番号、デバイス番号、サブチャンネル番号、そのボリュームへアクセスするときに利用可能なチャンネルバスID番号等が格納されている。またボリュームの状態としてボリュームが使用中か否か、つまり利用可能かどうかの情報も格納されている。

【0069】例えば、図6(a)のUCB17は、ボリューム名A13、デバイス番号100、アクセス可能状態、使用するチャンネルバスID CHPD01が格納されている状態を示している。デバイス番号とは、ボリューム毎に一意に付けられた番号である。

【0070】以上の情報構造により、DCB、DEB、UCBをたどり、利用可能なチャンネルバスIDがCHPD01であることを認識し、チャンネルバス9aおよびディスク制御装置8aを用いて、ボリュームA13をアクセスする。

【0071】これまでの説明から明らかなように、ホストとボリューム間のアクセス経路(チャンネルバス)は、SQA内のUCBの情報により制御されている。すなわち、UCB内の使用するチャンネルバスIDの情報を変更することにより、適用プログラムに影響を与えず(適用プログラムは実行中のまま)、利用するチャンネルバスを変更することが可能である。

【0072】図12(b)の状態に対応するようにチャンネルバスを再設定した、ホスト内部のディスクアクセスに関する制御構造を、図6(b)に示す。図6(a)と比較してわかるように、UCB17の使用するチャンネルバスIDの情報が、CHPD01からCHPD11に変更されている。この変更でボリュームA13へのアクセスは、このボリュームを管理するディスク制御装置8bへチャンネルバス9bを利用し直接行うことが可能となる。

【0073】ここで入出力制御ユーティリティとは、ディスク制御装置と、そのディスク制御装置に制御されるボリュームの対応の変更と同期して、UCB内の使用するチャンネルバスIDの情報を適切に変更するユーティリティプログラムである。これにより、ディスク制御装置に制御されるボリュームの対応の変更を、ホストコンピュータ上の適用プログラムから隠蔽できる。

【0074】入出力制御ユーティリティによる、ボリュームが利用するチャンネルバスの変更や入出力の保留の方式を図8を用いて説明する。図6で示した場合と同様、入出力制御ユーティリティの領域にも、入出力制御ユーテ

リティのUPA19と、入出力制御ユーティリティのSQA20が存在する。

【0075】ここでSQAは、全てのプログラムに対し共通領域となっている。すなわち、入出力制御ユーティリティのSQA20と、他のプログラムのSQA13は同じ領域である。これは、MVSオペレーティングシステムの機能として実現されている。このため、入出力制御ユーティリティはUCBの情報を書き換える事により、他のプログラムが利用している任意のボリュームの情報の変更が可能である。すなわち、チャンネルバスIDの情報を書き換えることにより利用するチャンネルバスの設定や変更を、ボリュームの状態の情報を「READY(利用可)」から「BUSY(使用中)」に書き換えることにより入出力の保留や再開の制御を行うことができる。

【0076】次に、本実施例における、ディスク制御装置間でのデータの移動を伴わないボリュームの再配置による負荷分散の具体的な手続きを、図7のフローチャートを用いて説明する。

【0077】各ディスク制御装置はそれぞれの資源に対する利用率をモニタし、各ディスク制御装置でモニタされたボリューム毎のディスク制御装置の負荷の情報は、デバイスネットワーク10等の情報交換網を通じてサービスプロセッサ11に集められる。この処理は、経路制御処理装置がサービスプロセッサに変わった以外は第一の実施例と同じである。

【0078】サービスプロセッサ11では、この負荷情報を元に各ディスク制御装置の負荷を計算し、限度以上の負荷の不均衡が生じている場合は、各ディスク制御装置の負荷が均等になるようなボリュームの配置を決定する。この限度は、あらかじめ利用者が指定しておく。具体的な配置は、自動最適配置が指定されている場合は、サービスプロセッサ11が各ディスク制御装置の負荷が均等になるように自律的に決定する。それ以外の場合は、ユーザが決定し指示する。また同様の機能を、サービスプロセッサではなくディスク制御装置に負わせることも同様に可能である。また、サービスプロセッサは、LAN(ローカルエリアネットワーク)等の通信回線18を経由して、ホストに接続されている。このため、サービスプロセッサはホストのコンソールや一般的な端末としての機能、すなわち、ホストのオペレーティングシステム等のパラメタの設定変更や、任意のホストプログラムの実行等の機能も持つ。

【0079】新しいボリューム配置の決定の後、サービスプロセッサはそのボリューム配置に最適になるような、アクセスパスの構成を計算する。ボリューム配置に最適になるようなアクセスパスの構成とは、図6(b)に示すように、そのボリュームに対するチャンネルバスIDの情報が、そのボリュームを制御するディスク制御装置につながっているチャンネルバスとなっている設定である。

【0080】計算された最適アクセスパス情報は、入出

力制御ユーティリティに送付される。入出力制御ユーティリティは、ホストシステム上で動作するユーティリティプログラムである。

【0081】最適アクセスパス情報を受け取った入出力制御ユーティリティは、その情報に従って、各ボリュームが利用するチャンネルパスの変更を行う。この変更在先立ち、チャンネルパスの変更を行うボリューム全てに対するアクセスを一時保留状態にする。保留状態への遷移は、各ボリュームのUCBのボリュームの状態情報を、使用中の状態にすることにより達成される。使用中のボリュームに対して発行されるI/Oアクセスは、オペレーティングシステムレベルで一時保留される。I/Oを発行した適用プログラムの立場では、単にI/Oの応答が戻ってくるまでの時間が延びたように見える。この使用中への移行は、チャンネルパス、ディスク制御装置やボリューム間の対応が変更されている間に、ボリュームへのアクセスが発生することを防ぎ、これら変更の手順を容易にするためである。

【0082】チャンネルパスの変更を行うボリューム全てに対するアクセスの一時保留状態への遷移が完了したら、入出力制御ユーティリティは、サービスプロセッサに入出力一時保留状態遷移完了報告を行う。サービスプロセッサはこの報告を受け、各ディスク制御装置にボリュームの再配置を指示する。ここでは第一の実施例と同様に、各ディスク制御装置は、ボリュームを移動し合う相手のディスク制御装置とボリューム情報等の交換を行ったのち、ボリュームの再配置を実行し、その完了をサービスプロセッサに報告する。

【0083】サービスプロセッサは、ボリューム移動の対象となった全てのディスク制御装置からの、再配置変更完了報告を確認の後、入出力制御ユーティリティに、各ボリュームが利用するチャンネルパスの変更の実行を指令する。

【0084】入出力制御ユーティリティは、各ボリュームが利用するチャンネルパスの変更を行った後、ボリュームに対するアクセスの一時保留を解除する。

【0085】これらの手順により、ホストコンピュータからディスク制御装置へのアクセスパス（チャンネルパス）のボリューム遷移と同期した再設定が実行できる。

【0086】第二の実施例においても、第一の実施例と同様に、デバイスネットワークで結合した任意台数のディスク制御装置と任意台数のディスク駆動装置及びボリュームに対して適用できる。また、ディスク制御装置の能力が足りない場合は、ディスク駆動装置の台数はそのまま、ディスク制御装置だけを増設し負荷分散することにより、システム全体のスループットを高めることもできる。

【0087】また、第一の実施例と同様に、ボリューム再配置を行っても実際のデータの移動が不要であり、秒オーダー以下の非常に短い時間での再配置が可能となる。また、ディスク制御装置間で共有し同時に参照し合う情

報が無い。すなわちボリューム再配置時以外は、ディスク制御装置とその配下のボリュームは、従来の単体のディスク制御装置と同様の動作を行う。特に、他のディスク制御装置との間で、ボリュームに格納する（されていた）入出力データのやりとりを行わない。これにより、ディスク制御装置間のスムーズな負荷分散の実現と、ディスク制御装置の増設によるスケーラブルな性能向上の両立が達成できる。

【0088】

10 【発明の効果】本発明によれば、複数のディスク制御装置間で、負荷分散となるボリュームの再配置を実際のデータ移動を伴わずに行い、従来型の単体のディスク制御装置複数個を用いた場合と比較して、より広いスケーラビリティと性能を得ることが出来る。またホストコンピュータがディスク制御装置およびボリュームにアクセスする時に用いるチャンネルパスを、前記ボリューム再配置に合わせ最適な経路となるように変更し、さらにその変更をホストコンピュータや適用プログラムから隠蔽することにより、適用プログラムやホストコンピュータの動作の停止が不要になり、前記ボリューム再配置による負荷分散の適用しやすさを高めることができる。

【図面の簡単な説明】

【図1】本発明における、ディスク制御装置とディスク駆動装置の一構成例である。

【図2】ディスク制御装置とディスク駆動装置の一構成例である。

【図3】ボリューム情報の一構成例を示す図である。

【図4】ディスク制御装置が、制御を担当するボリュームを変更することにより負荷分散を実現する方式の一つの例を示す図である。

【図5】ディスク制御装置が、制御を担当するボリュームを変更することにより負荷分散を実現するシステムの、ボリューム再配置の処理の流れを示す図である。

【図6】ホストコンピュータからディスク制御装置へのアクセスパスをボリューム遷移と同期して再設定する方式を示す図である。

【図7】ホストコンピュータからディスク制御装置へのアクセスパスをボリューム遷移と同期して再設定するシステムの、ボリューム再配置およびアクセスパス変更の処理の流れを示す図である。

【図8】ホストコンピュータからディスク制御装置へのアクセスパスを変更する方式を示す図である。

【図9】本発明における、ディスク制御装置とディスク駆動装置の第二の構成例である。

【図10】ホストコンピュータからディスク制御装置に渡すフレームの例を示す図である。

【図11】経路制御装置の一構成例を示す図である。

【図12】ディスク制御装置が、制御を担当するボリュームを変更することにより負荷分散を実現する方式の一つの例を示す図である。

19

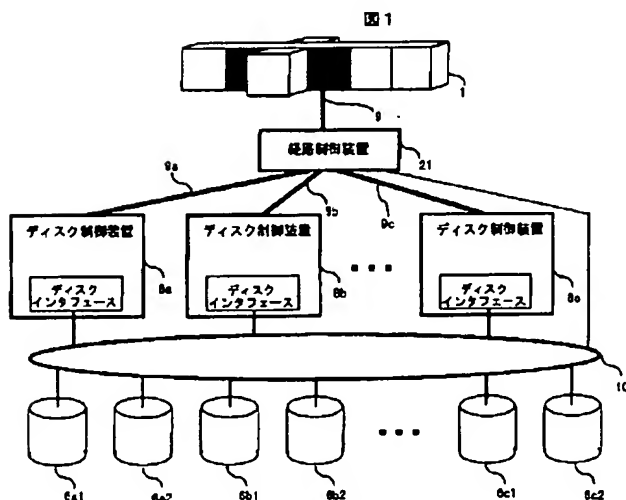
20

【符号の説明】

1…ホストコンピュータ、2…ホストインタフェース、
3…共有メモリ、4…キャッシュメモリ、5…ディスク
インタフェース、6・6a1・6a2・6b1・6b2
・6c1・6c2…ディスク駆動装置、7…共通バス、
8・8a・8b・8c…ディスク制御装置、9・9a
・9b・9c…チャネルバス、10…デバイスネットワ
ーク、11…サービスプロセッサ、12…ユーザプロ
グラムエリア、13…システムキューエリア、14…サブ
チャネル、15…データコントロールブロック、16…
データエクステンションブロック、17…ユニットコント
ロールブロック、18…LAN等の通信回線、19…入出力
制御ユーティリティのユーザプログラムエリア、20
…入出力制御ユーティリティのシステムキューエリア、

21…経路制御装置、22…送信先コントロールユニ
ット番号フィールド、23…発信元コントロールユニ
ット番号フィールド、24…ボリューム番号フィールド、25
…コマンド/データフィールド、26…ホストチャネル
インタフェース、27…ディスクチャネルインタフェ
ース、28…経路変更プロセッサ、29…ボリュームー
ディスク制御装置対応変換表、30…ボリューム再配置制
御プロセッサ、31…デバイスネットワークとのインタ
フェース、32…ボリューム情報、33…ボリューム番号フ
ィールド、34…ディスクインタフェース番号フ
ィールド、35…ディスク駆動装置番号フ
ィールド、36…HDA
番号フィールド、37…シリンダオフセットフ
ィールド。

【図1】

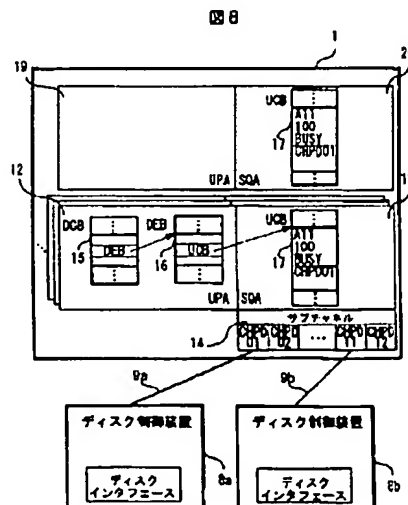


【図3】

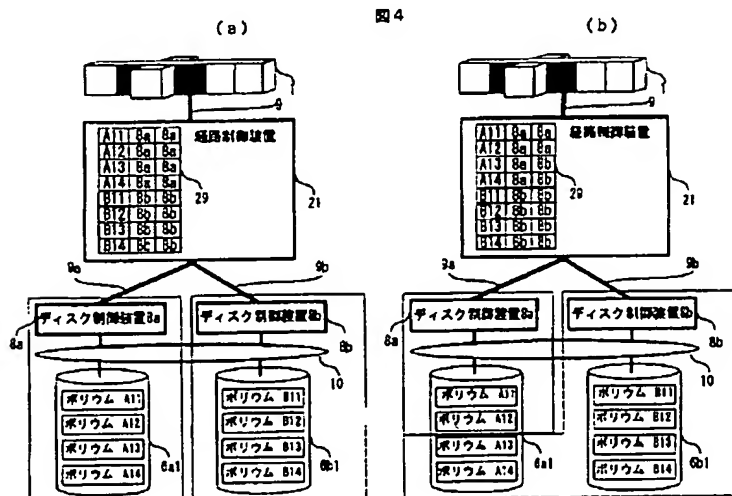
図 3

ボリューム番号	ディスク インタフェース 番号	ディスク 駆動装置 番号	HDA 番号	シリンダ オフセット
12	34	35	36	37
10	0	1	0	0
11	0	1	0	20
...

【図8】

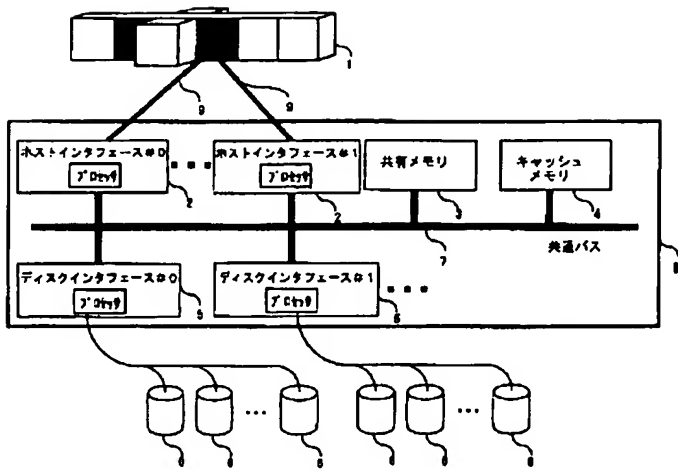


【図4】



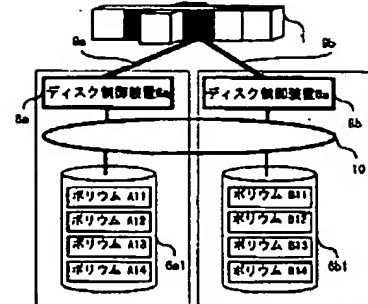
【図2】

図2

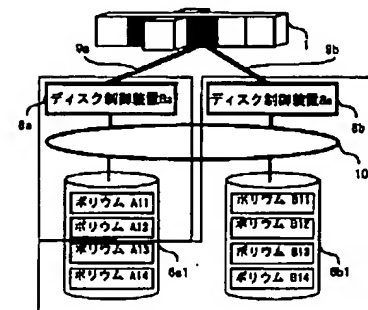


【図12】

図12 (a)

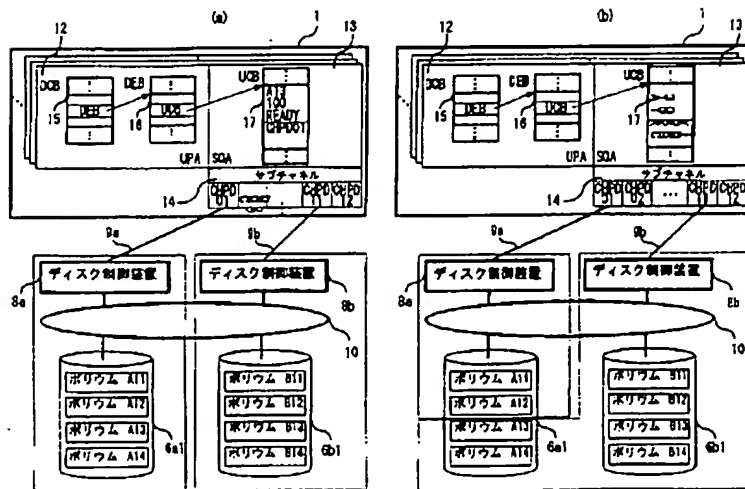


(b)



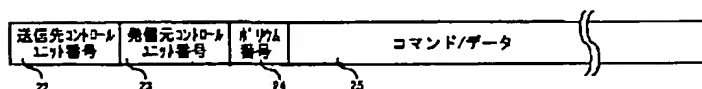
【図6】

図6



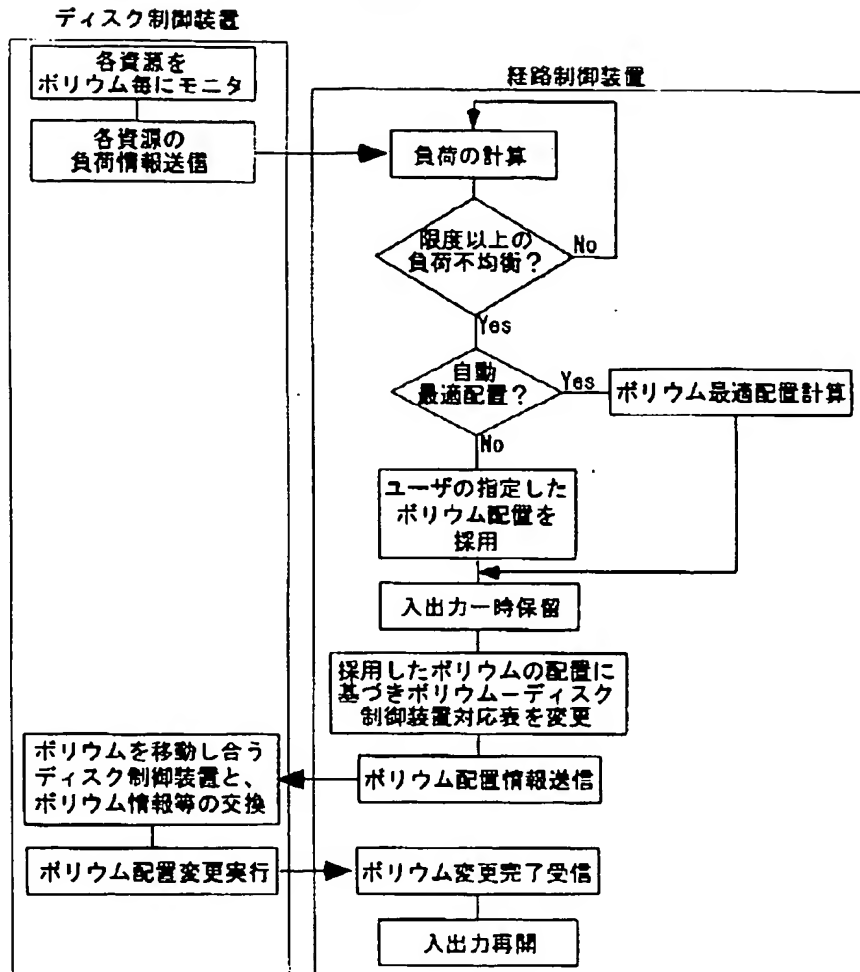
【図10】

図10



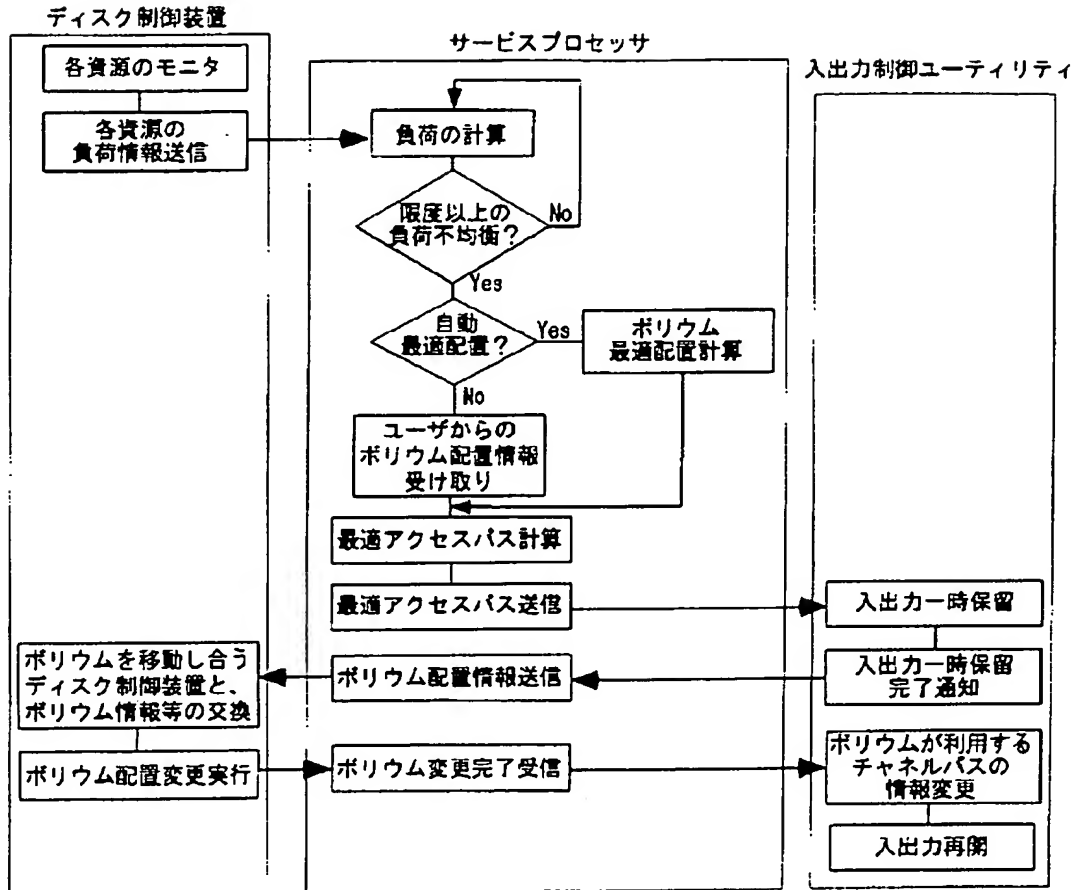
【図5】

図5



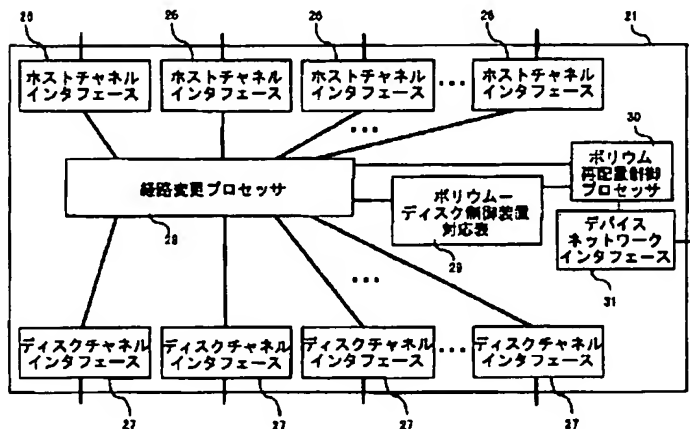
【図7】

図7



【図11】

図11



【図9】

図9

